A neural network for beam background decomposition of Belle II at SuperKEKB in real time

Yannik Buch, Ariane Frey, Lukas Herzberg, Benjamin Schwenker

Georg-August-Universität Göttingen
II. Physikalisches Institut

26.10.2023

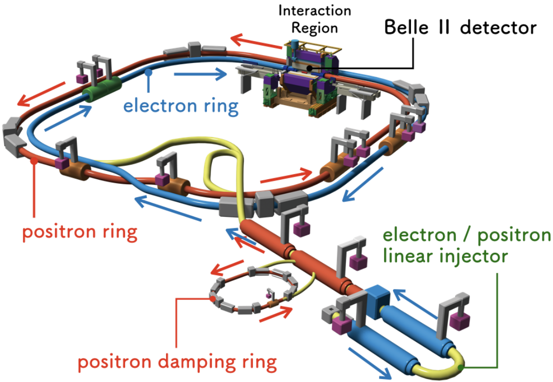
Federal Ministry of Education and Research

BG Net

GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN IN PUBLICA COMMODA SEIT 1737

Introduction: SuperKEKB

- Successor of KEKB, planned to take 50x the dataset of KEKB+Belle
- Asymmetric e^+e^- collider with $E_{\mathrm{CM}} = 10.48\,\mathrm{GeV}$
- B-factory precision machine at luminosity frontier to measure CPV
- Design Luminosity: $60 \times 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$
- World record (Juni 2022): $4.7 \times 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$
- Continuous top-up injections every $20\,\mathrm{ms}$ necessary due to short beam lifetime ($\sim 30\,\mathrm{min}$)
- Problems: Beam background too high and injections losses (dynamic aperture)



Introdution: Beam background

Storage

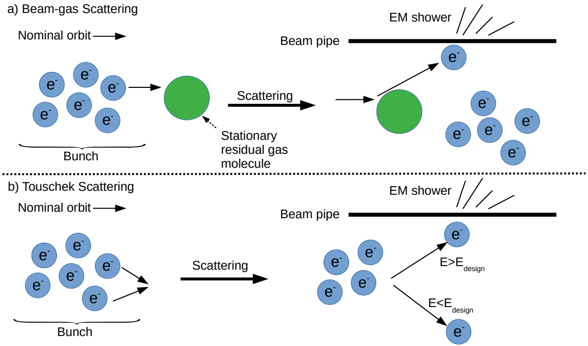
- Touschek scattering: Intra-bunch Coulomb scattering
- $\propto$ Current^2, 1/number bunches, 1/bunch volume
- Beam-gas: Interactions of particles with residual gas molecules
- $\propto$ Current, pressure in beam pipe

Lumi

- QED background: Radiative Bhabha
  Two-photon interactions
  $\rightarrow$ particles in Belle II
- $\propto$ instantaneous luminosity

Inj

- Noisy bunches: Betatron oscillations of injected bunches lead to high background
- Injections every 20 ms



Heuristics of beam backgrounds

- Interprete hit rates as product of heuristic G_1-G_8 and sensitivity S_1-S_8
- Heuristics describe known dependencies, e.g. beam pipe pressure, current
- Many state variables of unknown dependencies, e.g. collimator positions
- Neural networks parametrize the sensitivities/unknown dependencies
- The measured full hit rate $\mathcal{O}$ is obtained by adding all contributions

$$\mathcal{O} = \mathcal{O}_{\text{beam-gas, H}} + \mathcal{O}_{\text{Touschek, H}} + \mathcal{O}_{\text{beam-gas, L}} + \mathcal{O}_{\text{Touschek, L}} + \mathcal{O}_{\text{inj, H}} + \mathcal{O}_{\text{inj, L}} + \mathcal{O}_{\text{lumi}} + \mathcal{O}_{\text{ped}}$$

$$\mathcal{O}_{\text{beam-gas, (H/L)}} = S_{(1/3)} \times I_{(\text{H/L})} P_{(\text{H/L})} = S_{(1/3)} \times G_{(1/3)}$$

$$\mathcal{O}_{\text{Touschek, (H/L)}} = S_{(2/4)} \times \frac{I^2_{(\text{H/L})}}{n_{b,(\text{H/L})}\sigma_{x,(\text{H/L})}\sigma_{y,(\text{H/L})}\sigma_{z,(\text{H/L})}} = S_{(2/4)} \times G_{(2/4)}$$

$$\mathcal{O}_{\text{inj, H}} = S_5 \times G_5$$

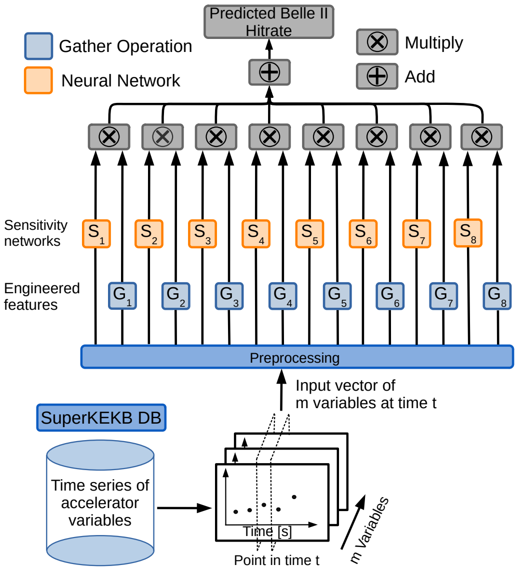
$$\mathcal{O}_{\text{inj, L}} = S_6 \times G_6$$

$$\mathcal{O}_{\text{lumi}} = S_7 \times \mathcal{L}$$

$$\mathcal{O}_{\text{ped}} = S_8$$

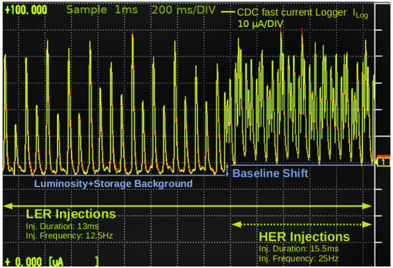
Neural network architecture

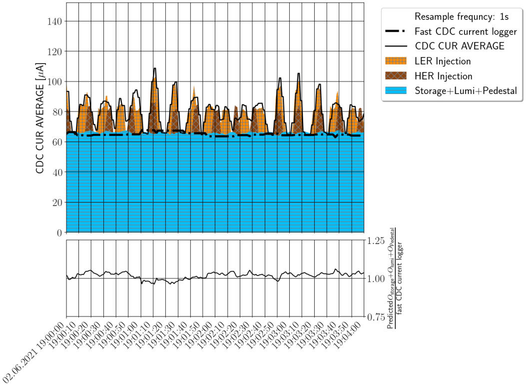
- Use 1 Hz time series of SuperKEKB variables
- Calculate analytical heuristics
- Use neural networks for sensitivities
- Neural networks are trained simultaneously
- Dense layers with tanh/softplus activation



Test of background decomposition

- Fast current logger of CDC records leakage current at 1000 Hz: Can resolve injections
- Need to resample fast current logger to compare with BGNet prediction
- Injection background falls to zero between injections: fit baseline through minima
- Separate storage+lumi+ped and injection backgrounds for CDC



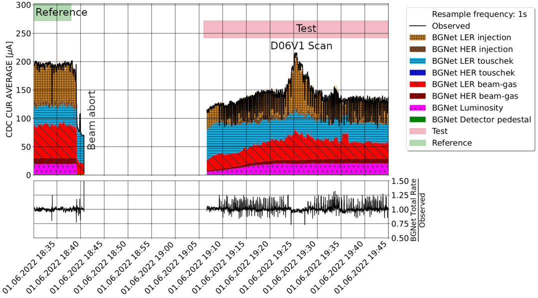


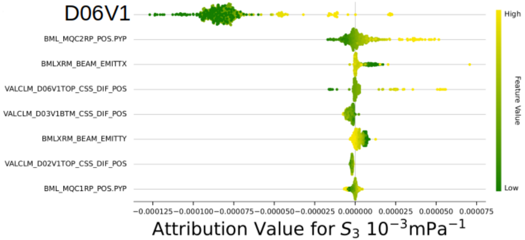
Feature attribution with expected gradients

- Feature attributions quantify the impact on the model prediction of a change of input values
- The change of input values is measured between a reference and a test set

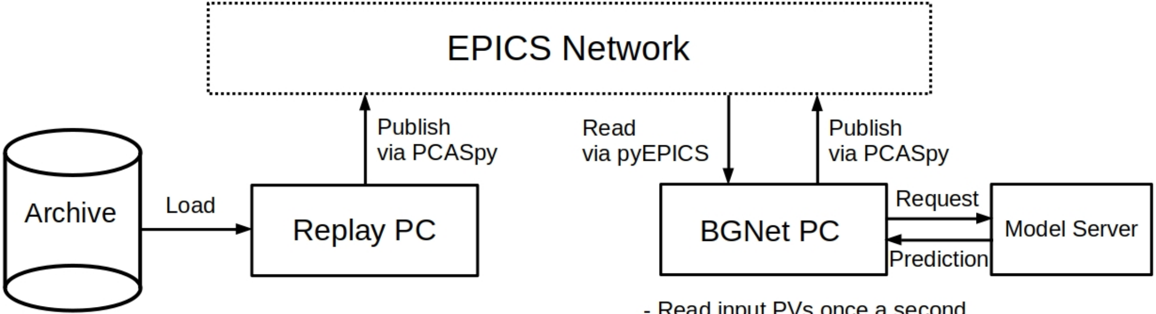
$\Rightarrow$ If the model follows a change in measured hit rates then the responsible input(s) can be identified

- Example: A collimator scan of D06V1 after a beam abort that damaged the collimator





Test of the real time application



- Archive holds historic data from 2022 that is replayed
- Publish all relevant PVs once every second

- Read input PVs once a second
- Send request to model server
- Publish predictions to EPICS network

Display - Attributions

Rank	PV	Attributions	Gradient
1	SKB_HER_P_avg	6.79400571e-08	-
2	BMHXRM_BEAM_EMITTY	7.86373e-09	-
3	BMHXRM_BEAM_EMITTX	4.68460781e-09	+
4	VAHCLM_D01V1BTM_CSS_DIF_POS	8.04621203e-10	-
5	VAHCLM_D01V1TOP_CSS_DIF_POS	1.70636574e-10	+

Show attribution values for 5 PVs with highest attribution

Ranking of PVs with largest attribution → likely candidates causing change in hit rate

Show the sensitivity of background component and the value of PV

Option to refresh at any point to re-evaluate ranking (= PVs with highest attribution)

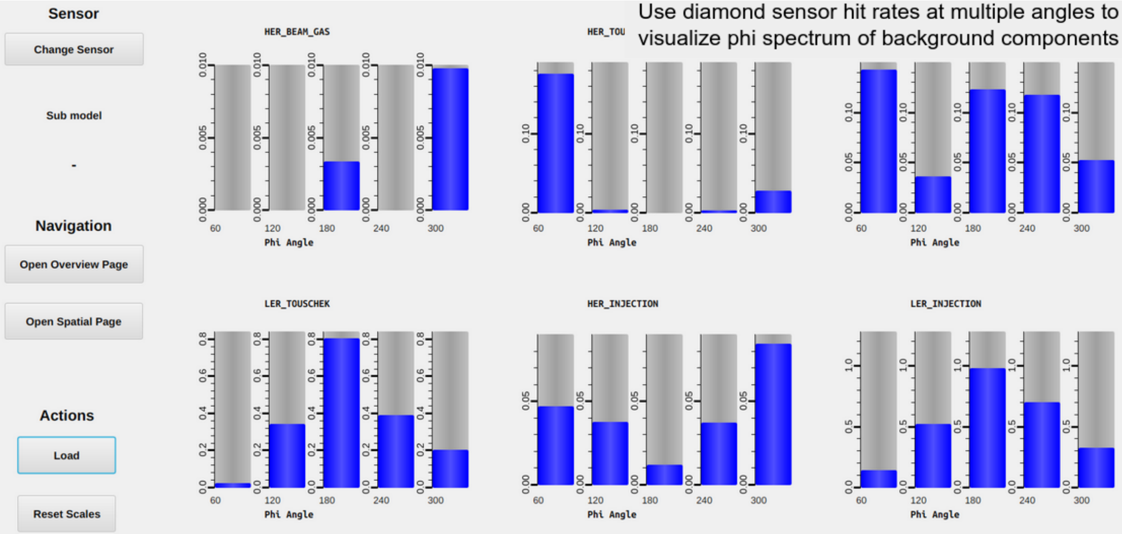
Summary and Outlook

- A neural network for the decomposition of beam backgrounds was implemented
- The results are published in NimA BGNet Paper
- Feature attribution methods can be used to identify important variables
- Demonstrator for the real time application using EPICS+BGNet was succesful with latencies <1s
- CS-Studio displays were creared to communicate BGNet results to machine operators
- Setup is installed and integrated at KEK and tested once long shutdown 1 is over (beginning 2024)
- Implement technical and visual improvements

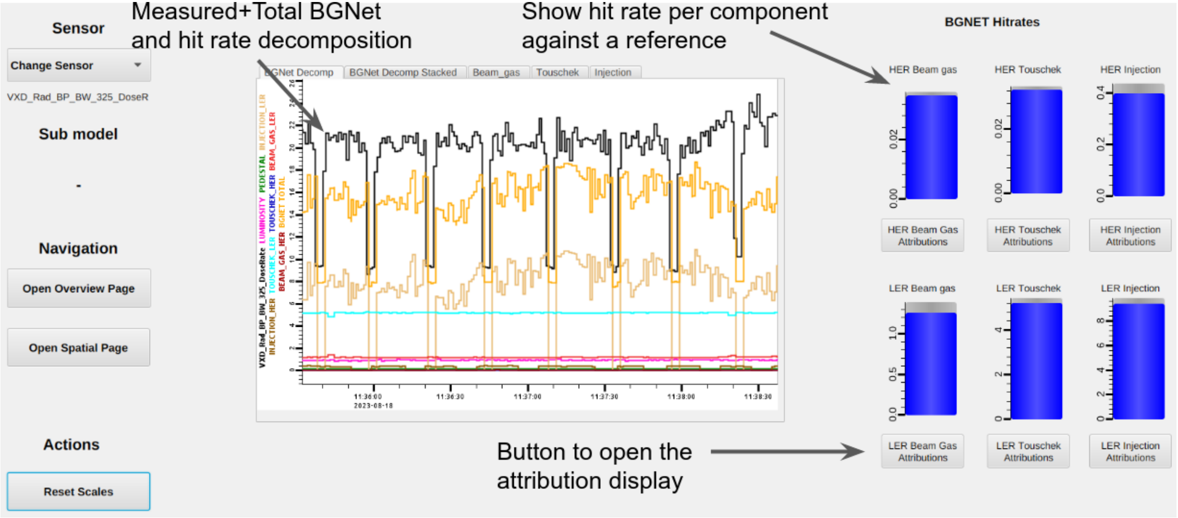
Backup

Display - Spatial

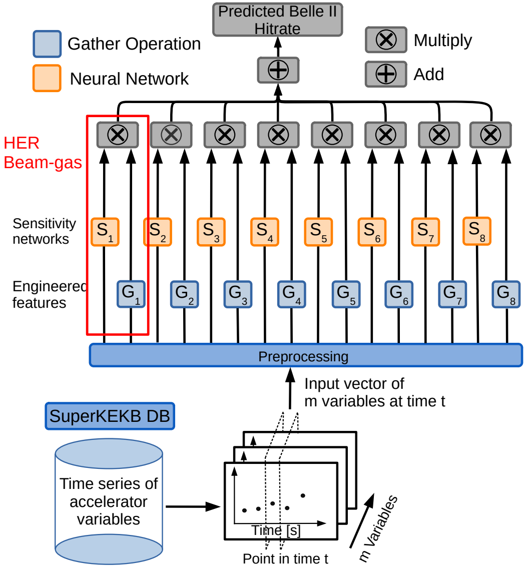
Use diamond sensor hit rates at multiple angles to visualize phi spectrum of background components



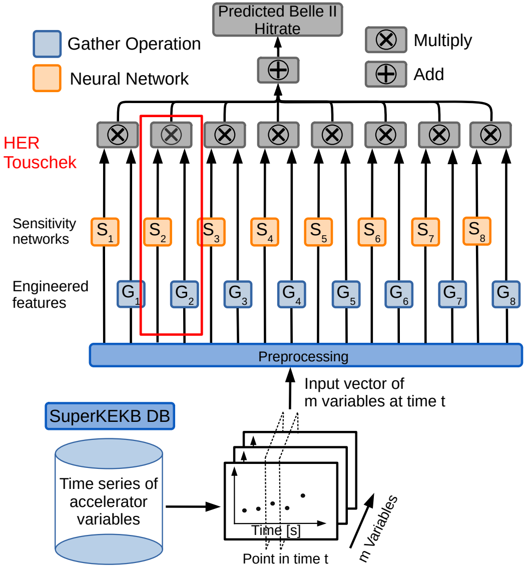
Display - Overview



Neuronales Netzwerk Aufbau

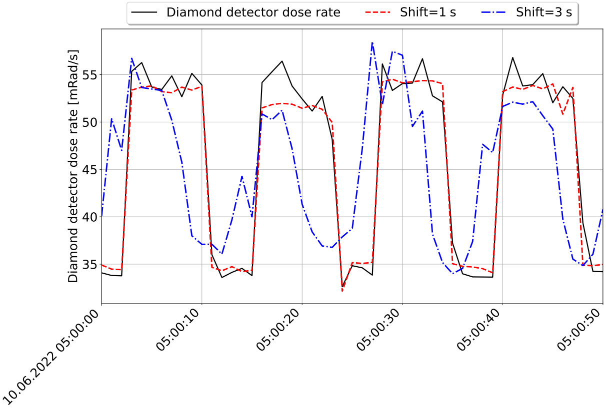


Aufbau neuronales Netzwerk



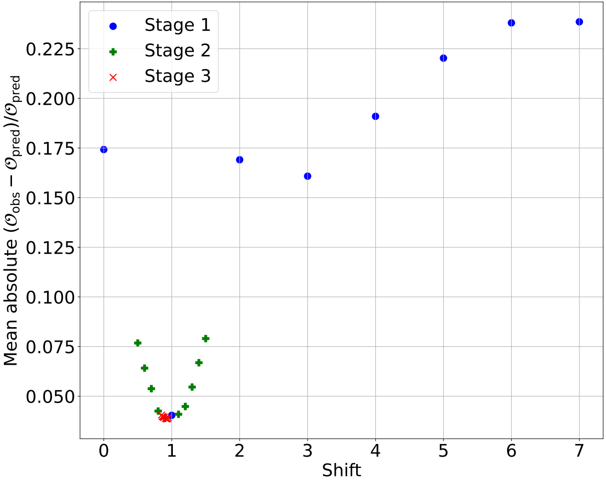
Delay correction

- Data from EPICS slow control system is not synchronized
- Use weighted average of two data points to correct
- Repeat training for different delays and use training with smallest loss



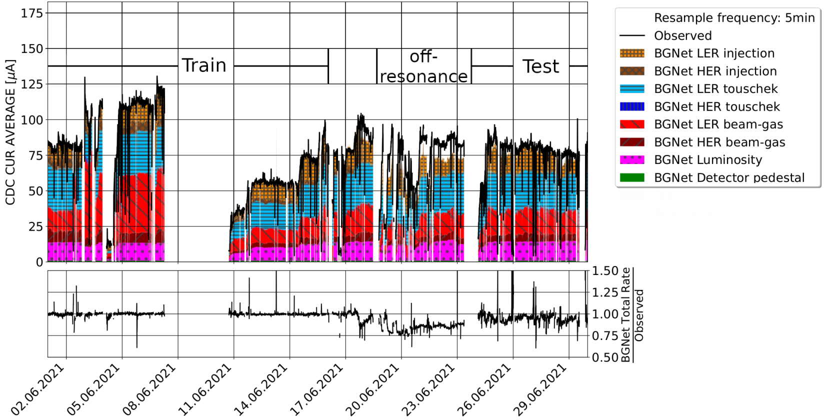
Delay correction

- Data from EPICS slow control system is not synchronized
- Use weighted average of two data points to correct
- Repeat training for different delays and use training with smallest loss



Extrapolation

- Model extrapolation using data during June 2021
- Off-resonance referes to data taking with energies other than $\Upsilon(4S)$ resonance



Summary

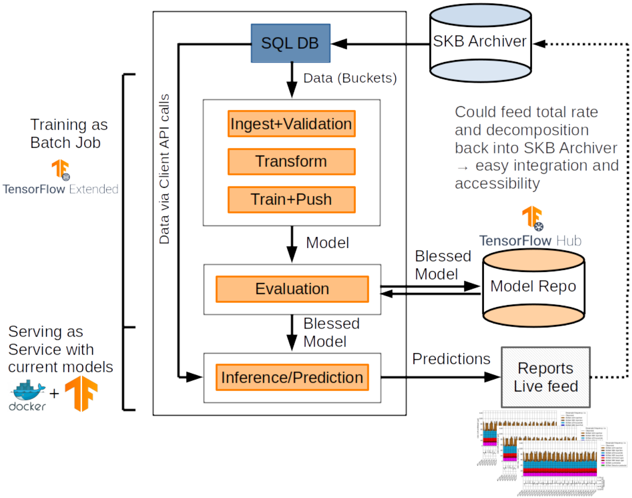
- Built model using heuristic scaling laws
- Explore the possible insights into the background composition BGNet can provide
- For more information see our paper:

  Title: A neural network for beam background decomposition in Belle II at SuperKEKB
  NimA: Volume 1049, April 2023, 168112

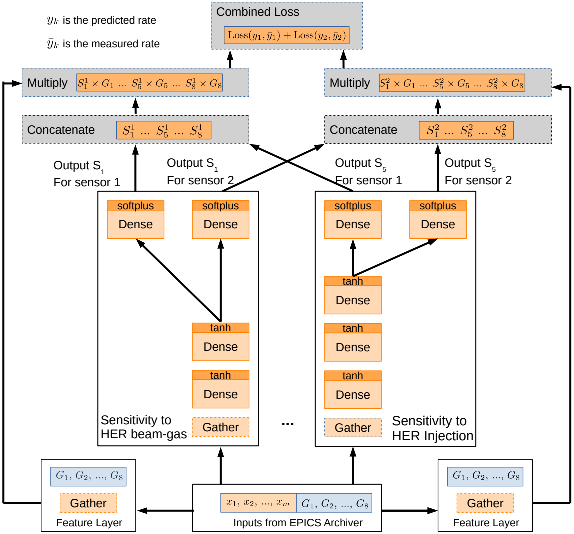
Going forward:

- Exploit scaling properties of background to use one network for multiple sub-detectors
- Implement real-time display for use by SuperKEKB operators
- Help control room operators to optimize background conditions

real-time alt

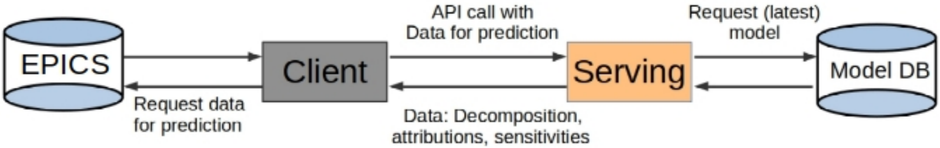


Multi detector network



Real-time monitoring

- Set up BGNet decomposition as live-feed accessible in KEK control room
- Obtain live values from EPICS network
- Do inference with tensorflow model server
- Ingest decomposition back into EPICS network



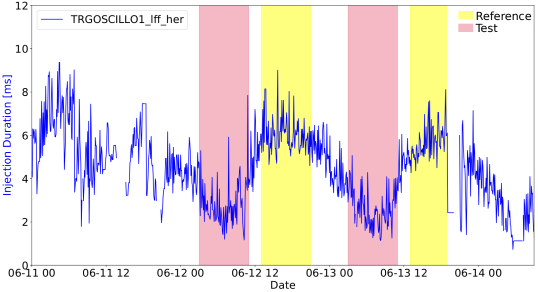
Explain backgrounds

- Use feature attribution to explain oscillations in injection duration
- Path Explain: arXiv:2002.04138, https://github.com/suinleelab/path_explain
- Test set x, reference set x' $\rightarrow$ compare $f(x)$ to $f(x')$, $y = f(x)$ functional representation of NN
- Then ask how much individual inputs x_i (i-th input variable) contribute to difference $f(x) - f(x')$

$$\boxed{\text{Attributions}_i(x) = \int_{x'} \left((x_i - x'_i) \times \int_{\alpha=0}^{1} \frac{\delta f(x' + \alpha(x - x'))}{\delta x_i} d\alpha \right) p_D(x') dx'}$$

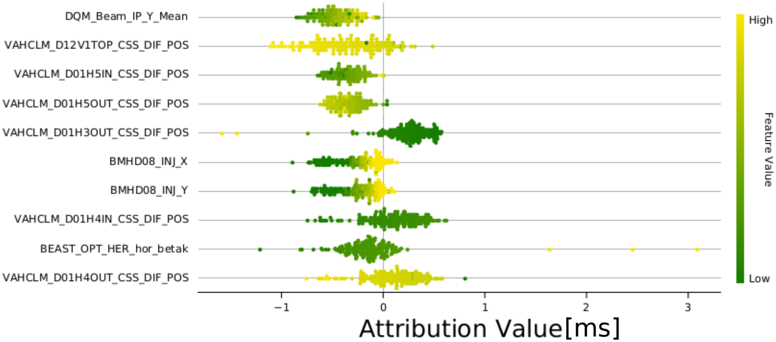
$$\int_{\alpha=0}^{1} \frac{\delta f(x' + \alpha(x - x'))}{\delta x_i} d\alpha$$

- Instead of calc. attribution at x or x', calc. along shortest path
- **Retains completeness**
- Entire expression model agnostic

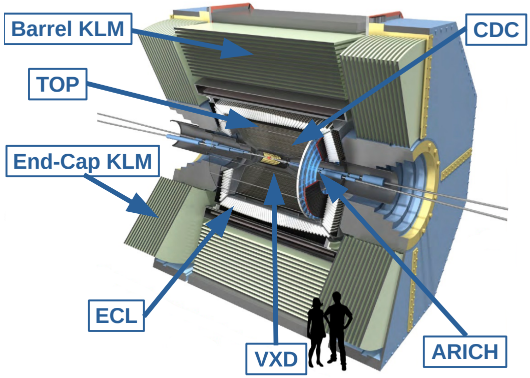


Explain backgrounds

- Mean absolute value of the attribution per input feature ranks the variables
- The color gradient describes the relative numeric value of the input variable
- Explaining which inputs changed from minimum to maximum



Introduction: Belle II



- Asymmetric 4π-coverage detector
- 2 ns bunch spacing
- Tracking: VXD (pixel and double-sided strip detector) + CDC
- Particle ID: TOP and ARICH
- ECL delivers energy and position measurement
- K_L and muon detector in barrel and end-cap region
- Radiation monitoring with single-crystal diamond sensors

Neural network for background decomposition

- Use heuristic scaling laws G_i multiplied with unknown sensitivity S_i
- Regression target: Belle II detector hitrates $\mathcal{O}$
- Training data: SuperKEKB variables describing machine state

$$\mathcal{O} = \mathcal{O}_{\text{beam-gas, H}} + \mathcal{O}_{\text{Touschek, H}} + \mathcal{O}_{\text{beam-gas, L}} + \mathcal{O}_{\text{Touschek, L}} + \mathcal{O}_{\text{inj, H}} + \mathcal{O}_{\text{inj, L}} + \mathcal{O}_{\text{lumi}} + \mathcal{O}_{\text{ped}} \tag{1}$$

$$\mathcal{O}_{\text{beam-gas, (H/L)}} = S_{(1/3)} \times I_{(\text{H/L})} P_{(\text{H/L})} = S_{(1/3)} \times G_{(1/3)} \tag{2}$$

$$\mathcal{O}_{\text{Touschek, (H/L)}} = S_{(2/4)} \times \frac{I^2_{(\text{H/L})}}{n_{b,(\text{H/L})} \sigma_{x,(\text{H/L})} \sigma_{y,(\text{H/L})} \sigma_{z,(\text{H/L})}} = S_{(2/4)} \times G_{(2/4)} \tag{3}$$

$$\mathcal{O}_{\text{inj, H}} = S_5 \times G_5 \tag{4}$$

$$\mathcal{O}_{\text{inj, L}} = S_6 \times G_6 \tag{5}$$

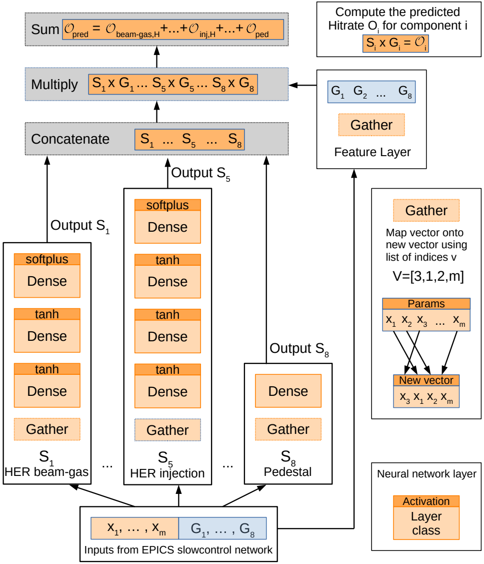
$$\mathcal{O}_{\text{lumi}} = S_7 \times \mathcal{L} \tag{6}$$

$$\mathcal{O}_{\text{ped}} = S_8 \tag{7}$$

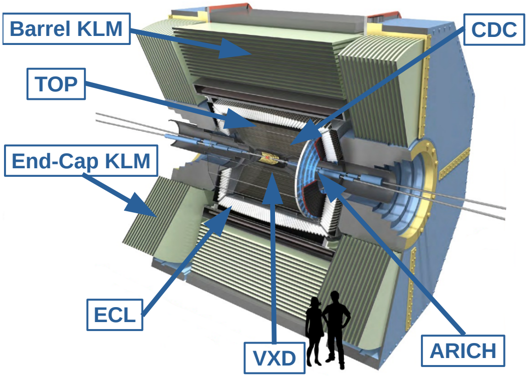
- σ are bunch sizes, $\mathcal{L}$ is the inst. lumi.
- G_5 and G_6 are 1 during top-up injections and 0 otherwise

Neural network for background decomposition

- Get selected subset of SuperKEKB variables
- Calculate heuristic features G_i
- Training loss is MAE of predicted and measured hit rate
- Outputs of last three layers are available: sensitivities, background components, total predicted hitrate



Einleitung: Belle II



- Asymmetrischer 4π Detektor
- 2 ns bunch spacing
- Tracking: VXD (pixel und double-sided strip Detektor) + CDC
- Particle ID: TOP und ARICH
- ECL liefert Energie- und Positionsmessung
- K_L und Myon Detektor in der barrel und end-cap Region